

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/568,396  
Source: IFWP  
Date Processed by STIC: 2/27/06

# ***ENTERED***



IFWP

## RAW SEQUENCE LISTING

DATE: 02/27/2006

PATENT APPLICATION: US/10/568,396

TIME: 14:54:36

Input Set : A:\67489-PCT-US\_SeqList.txt

Output Set: N:\CRF4\02272006\J568396.raw

```

4 <110> APPLICANT: Stephen P. Goff
5   Guanxia Gao
7 <120> TITLE OF INVENTION: ZAP PROTEIN AND RELATED COMPOSITIONS AND METHODS
9 <130> FILE REFERENCE: 67489-PCT-US/JPW/JW
C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/568,396
C--> 11 <141> CURRENT FILING DATE: 2006-02-13
11 <150> PRIOR APPLICATION NUMBER: PCT/US2004/026162
12 <151> PRIOR FILING DATE: 2004-08-12
14 <160> NUMBER OF SEQ ID NOS: 10
16 <170> SOFTWARE: PatentIn version 3.1
18 <210> SEQ ID NO: 1
19 <211> LENGTH: 776
20 <212> TYPE: PRT
21 <213> ORGANISM: mammalian
23 <400> SEQUENCE: 1
25 Met Ala Asp Pro Gly Val Cys Cys Phe Ile Thr Lys Ile Leu Cys Ala
26 1           5           10           15
28 His Gly Gly Arg Met Thr Leu Glu Glu Leu Leu Gly Glu Ile Arg Leu
29           20           25           30
31 Pro Glu Ala Gln Leu Tyr Glu Leu Leu Glu Thr Ala Gly Pro Asp Arg
32           35           40           45
34 Phe Val Leu Leu Glu Thr Gly Gly Gln Ala Gly Ile Thr Arg Ser Val
35           50           55           60
37 Val Ala Thr Thr Arg Ala Arg Val Cys Arg Arg Lys Tyr Cys Gln Arg
38 65           70           75           80
40 Pro Cys Asp Ser Leu His Leu Cys Lys Leu Asn Leu Leu Gly Arg Cys
41           85           90           95
43 His Tyr Ala Gln Ser Gln Arg Asn Leu Cys Lys Tyr Ser His Asp Val
44           100          105          110
46 Leu Ser Glu Gln Asn Phe Gln Ile Leu Lys Asn His Glu Leu Ser Gly
47           115          120          125
49 Leu Asn Gln Glu Glu Leu Ala Cys Leu Leu Val Gln Ser Asp Pro Phe
50           130          135          140
52 Phe Leu Pro Glu Ile Cys Lys Ser Tyr Lys Gly Glu Gly Arg Lys Gln
53 145          150          155          160
55 Thr Cys Gly Gln Pro Gln Pro Cys Glu Arg Leu His Ile Cys Glu His
56           165          170          175
58 Phe Thr Arg Gly Asn Cys Ser Tyr Leu Asn Cys Leu Arg Ser His Asn
59           180          185          190
61 Leu Met Asp Arg Lys Val Leu Thr Ile Met Arg Glu His Gly Leu Ser
62           195          200          205
64 Pro Asp Val Val Gln Asn Ile Gln Asp Ile Cys Asn Asn Lys His Ala
65           210          215          220

```

## RAW SEQUENCE LISTING

DATE: 02/27/2006

PATENT APPLICATION: US/10/568,396

TIME: 14:54:36

Input Set : A:\67489-PCT-US\_SeqList.txt

Output Set: N:\CRF4\02272006\J568396.raw

```

67 Arg Arg Asn Pro Pro Gly Thr Arg Ala Ala His Pro His Arg Arg Gly
68 225                230                235                240
70 Gly Ala His Arg Asp Arg Ser Lys Ser Arg Asp Arg Phe Leu His Asn
71                245                250                255
73 Ser Leu Glu Phe Leu Ser Pro Val Val Ser Pro Leu Gly Ser Gly Pro
74                260                265                270
76 Pro Ser Pro Asp Val Thr Ser Cys Lys Asp Ser Leu Glu Asp Val Ser
77                275                280                285
79 Val Asp Val Thr Gln Lys Phe Lys Tyr Leu Gly Thr His Asp Arg Ala
80                290                295                300
82 Gln Leu Ser Pro Val Ser Ser Lys Ala Ala Gly Val Gln Gly Pro Ser
83 305                310                315                320
85 Gln Met Arg Ala Ser Gln Glu Phe Ser Glu Asp Gly Asn Leu Asp Asp
86                325                330                335
88 Ile Phe Ser Arg Asn Arg Ser Asp Ser Ser Ser Arg Ala Ser Ala
89                340                345                350
91 Ala Lys Val Ala Gln Arg Asn Glu Ala Val Ala Met Lys Met Gly Met
92                355                360                365
94 Glu Val Lys Gly Lys Lys Glu Ala Pro Asp Ile Asp Arg Val Pro Phe
95                370                375                380
97 Leu Asn Ser Tyr Ile Asp Gly Val Thr Met Glu Lys Ala Ser Val Ser
98 385                390                395                400
100 Gly Ile Pro Gly Lys Lys Phe Thr Ala Asn Asp Leu Glu Asn Leu Leu
101                405                410                415
103 Leu Leu Asn Asp Thr Trp Lys Asn Val Ala Lys Pro Gln Asp Leu Gln
104                420                425                430
106 Thr Thr Gly Arg Ile Thr Asp Ser Gly Gln Asp Lys Ala Phe Leu Gln
107                435                440                445
109 Asn Lys Tyr Gly Gly Asn Pro Val Trp Ala Ser Ala Ser Thr His Asn
110                450                455                460
112 Ala Pro Asn Gly Ser Ser Gln Ile Met Asp Glu Thr Pro Asn Val Ser
113 465                470                475                480
115 Lys Ser Ser Thr Ser Gly Phe Ala Ile Lys Pro Ala Ile Ala Gly Gly
116                485                490                495
118 Lys Glu Ala Val Tyr Ser Gly Val Gln Ser Pro Arg Ser Gln Val Leu
119                500                505                510
121 Ala Val Pro Gly Glu Ala Thr Thr Pro Val Gln Ser Asn Arg Leu Pro
122                515                520                525
124 Gln Ser Pro Leu Ser Ser Ser Ser His Arg Ala Ala Ala Ser Gly Ser
125                530                535                540
127 Pro Gly Lys Asn Ser Thr His Thr Ser Val Ser Pro Ala Ile Glu Ser
128 545                550                555                560
130 Ser Arg Met Thr Ser Asp Pro Asp Glu Tyr Leu Leu Arg Tyr Ile Leu
131                565                570                575
133 Asn Pro Leu Phe Arg Met Asp Asn His Gly Pro Lys Glu Ile Cys Gln
134                580                585                590
136 Asp His Leu Tyr Lys Gly Cys Gln Gln Ser His Cys Asp Arg Ser His
137                595                600                605
139 Phe His Leu Pro Tyr Arg Trp Gln Met Phe Val Tyr Thr Thr Trp Arg

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/568,396

DATE: 02/27/2006

TIME: 14:54:36

Input Set : A:\67489-PCT-US\_SeqList.txt

Output Set: N:\CRF4\02272006\J568396.raw

```

140      610      615      620
142 Asp Phe Gln Asp Met Glu Ser Ile Glu Gln Ala Tyr Cys Asp Pro His
143 625      630      635      640
145 Val Glu Leu Ile Leu Ile Glu Asn His Gln Ile Asn Phe Gln Lys Met
146      645      650      655
148 Thr Cys Asp Ser Tyr Pro Ile Arg Arg Leu Ser Thr Pro Ser Tyr Glu
149      660      665      670
151 Glu Lys Pro Leu Ser Ala Val Phe Ala Thr Lys Trp Ile Trp Tyr Trp
152      675      680      685
154 Lys Asn Glu Phe Asn Glu Tyr Ile Gln Tyr Gly Asn Glu Ser Pro Gly
155      690      695      700
157 His Thr Ser Ser Asp Ile Asn Ser Ala Tyr Leu Glu Ser Phe Phe Gln
158 705      710      715      720
160 Ser Cys Pro Arg Gly Val Leu Pro Phe Gln Ala Gly Ser Gln Lys Tyr
161      725      730      735
163 Glu Leu Ser Phe Gln Gly Met Ile Gln Thr Asn Ile Ala Ser Lys Thr
164      740      745      750
166 Gln Arg His Val Val Arg Arg Pro Val Phe Val Ser Ser Asn Asp Val
167      755      760      765
169 Glu Gln Lys Arg Arg Gly Pro Glu
170      770      775
173 <210> SEQ ID NO: 2
174 <211> LENGTH: 2331
175 <212> TYPE: DNA
176 <213> ORGANISM: mammalian
178 <400> SEQUENCE: 2
179 atggcagatc cgggggtatg ctgtttcatc accaagatcc tgtgcgcca cgggggcccgt      60
181 atgaccctgg aggaactgct ggggtgagatc aggcctccccg aggcgcagct ctacgagctg      120
183 ctggagacgg cggggcccga tcgcttcgtg ctattggaga ctggaggcca ggccgggcatc      180
185 actcgggtctg tagtggctac tactcgagcc cgcgtctgcc gtcggaagta ctgccagaga      240
187 ccctgcgaca gcctgcacct ctgcaagctt aatctgctcg gccggtgcca ctatgcacag      300
189 tctcagcgga acctctgcaa atattctcac gatgttctct cggaacagaa cttccagatc      360
191 ctgaagaatc atgagctctc tgggcttaac caagaggagc tagcttgctt cctgggtccaa      420
193 agcgaccctt ttttcttgcc cgagatatgc aagagttaca aaggagaggg ccgaaaacag      480
195 acctgtgggc agccacagcc atgcgagaga ctccacatct gtgagcactt caccgggggc      540
197 aactgcagtt acctcaactg tctcaggtct cacaacctga tggacagaaa ggtgttgacc      600
199 atcatgaggg agcacgggct ggtcctgat gtggtccaga acatccagga catctgcaac      660
201 aacaaacacg ccaggaggaa cccgcctggc acgagagctg cccatccaca ccgcagaggc      720
203 ggcgcacaca gagacagaag caaaagcaga gaccgcttcc ttcacaacag tctagaatct      780
205 ctctcacctg ttgtctcacc tctgggatct ggtccgcta gccagatgt caccagctgt      840
207 aaagattccc tggaggatgt gtctgtggat gtcaccaga agttcaagta cttggggacg      900
209 catgaccgtg cgcagctctc cccagtctca tctaaggctg ctggtgttca aggaccagt      960
211 caaatgagag caagccaaga gttttcagag gatgggaatc tagatgacat attttctagg      1020
213 aatcgttctg attcatcatc aagtcgagcc tccgctgcca aggtggcaca aagaaatgaa      1080
215 gctgtggcca tgaaaatggg catggaggtc aagggcaaga aggaggctcc agacatcgat      1140
217 cgggtcccat ttttaaatag ttatattgat ggggtgacca tggaaaaagc atcgggtctca      1200
219 ggaattccag gcaaaaagtt cacagccaat gatctggaaa atttgctatt acttaacgac      1260
221 acttgggaaga atgtggctaa gcccaggat ctgcagacca caggcagaat cactgacagt      1320
223 ggccaagaca aggcattcct gcagaataaa tatggaggaa acccagtggtg ggcaagtgca      1380

```

## RAW SEQUENCE LISTING

DATE: 02/27/2006

PATENT APPLICATION: US/10/568,396

TIME: 14:54:36

Input Set : A:\67489-PCT-US\_SeqList.txt

Output Set: N:\CRF4\02272006\J568396.raw

```

225 tccacccata atgccccaaa tggctctagt caaattatgg atgaaactcc taatgtctct 1440
227 aaaagtagta ccagtggttt tgccataaaa ccagcaattg ctggaggaaa agaagcagtc 1500
229 tattctggag ttcagagtc gagaaagccag gtccctagctg tgccctgggga ggctactacc 1560
231 cctgtacaga gcaacaggct gcctcagtcg cctctgtctt cctcaagcca cagagctgca 1620
233 gcctctggga gccctggcaa gaactccacc catacctctg tgagcccagc catcgagtct 1680
235 tcaaggatga catcagaccc cgatgagtat ctccctacgt acatcctaaa tcctttatct 1740
237 aggatggata atcatggccc gaaggaaatc tgtcaggacc atctgtacaa gggctgtcaa 1800
239 cagagccact gcgacaggag tcaattccat ctgccctacc ggtggcagat gttcgtatat 1860
241 accacttgga gggacttcca ggacatggag tctatcgaa aggcctattg tgatccccac 1920
243 gttgaactca ttttgataga aaaccatcag atcaatttcc agaaaatgac ctgtgactcc 1980
245 taccatccac gacgcctctc cactccctca tatgaggaaa agccacttag tgctgtcttc 2040
247 gccaccaagt ggatttggtt ttggaagaat gaatttaatg aatatatcca gtatgggaat 2100
249 gagagcccag gccacaccag ctctgacatc aactctgcgt acctggagtc tttcttccag 2160
251 tcttgtccca ggggagtttt gccattccag gctgggtcac agaagtacga gttaagcttc 2220
253 caagggatga ttcagacaaa tatagcttcc aagactcaaa ggcattgtgt cagaaggcca 2280
255 gtatttgttt cttcgaacga tgtggagcag aagagaagag gtccagagtg a 2331
258 <210> SEQ ID NO: 3
259 <211> LENGTH: 78
260 <212> TYPE: DNA
261 <213> ORGANISM: Artificial Sequence
263 <220> FEATURE:
264 <223> OTHER INFORMATION: PCR Primer
266 <400> SEQUENCE: 3
267 ataagcttgc caccatggct tctccststg stgttccaga tatgctgaat tcggcgcccg 60
269 cgccaagttg accagtgc 78
272 <210> SEQ ID NO: 4
273 <211> LENGTH: 26
274 <212> TYPE: DNA
275 <213> ORGANISM: Artificial Sequence
277 <220> FEATURE:
278 <223> OTHER INFORMATION: PCR Primer
280 <400> SEQUENCE: 4
281 atatcgattc agtctgctc ctggc 26
284 <210> SEQ ID NO: 5
285 <211> LENGTH: 38
286 <212> TYPE: DNA
287 <213> ORGANISM: Artificial Sequence
289 <220> FEATURE:
290 <223> OTHER INFORMATION: Oligonucleotide
292 <400> SEQUENCE: 5
293 ctagataact tcgtataatg tatgctatac gaagttat 38
296 <210> SEQ ID NO: 6
297 <211> LENGTH: 38
298 <212> TYPE: DNA
299 <213> ORGANISM: Artificial Sequence
301 <220> FEATURE:
302 <223> OTHER INFORMATION: Oligonucleotide
304 <400> SEQUENCE: 6
305 ctagataact tcgtatagca tacattatac gaagttat 38

```

## RAW SEQUENCE LISTING

DATE: 02/27/2006

PATENT APPLICATION: US/10/568,396

TIME: 14:54:36

Input Set : A:\67489-PCT-US\_SeqList.txt

Output Set: N:\CRF4\02272006\J568396.raw

```

308 <210> SEQ ID NO: 7
309 <211> LENGTH: 25
310 <212> TYPE: DNA
311 <213> ORGANISM: Artificial Sequence
313 <220> FEATURE:
314 <223> OTHER INFORMATION: PCR Primer
316 <400> SEQUENCE: 7
317 gcttatccat atgatgttcc agatt                25
320 <210> SEQ ID NO: 8
321 <211> LENGTH: 34
322 <212> TYPE: DNA
323 <213> ORGANISM: Artificial Sequence
325 <220> FEATURE:
326 <223> OTHER INFORMATION: PCR Primer
328 <400> SEQUENCE: 8
329 atataggcgg ccgccctctg gacctcttct cttc      34
332 <210> SEQ ID NO: 9
333 <211> LENGTH: 19
334 <212> TYPE: DNA
335 <213> ORGANISM: Artificial Sequence
337 <220> FEATURE:
338 <223> OTHER INFORMATION: PCR Primer
340 <400> SEQUENCE: 9
341 gagctctctg ggcttaacc                19
344 <210> SEQ ID NO: 10
345 <211> LENGTH: 34
346 <212> TYPE: DNA
347 <213> ORGANISM: Artificial Sequence
349 <220> FEATURE:
351 <223> OTHER INFORMATION: PCR Primer
353 <400> SEQUENCE: 10
354 atataggcgg ccgccctctg gacctcttct cttc      34

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/568,396

DATE: 02/27/2006

TIME: 14:54:37

Input Set : A:\67489-PCT-US\_SeqList.txt

Output Set: N:\CRF4\02272006\J568396.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date